LabeledIn: Cataloging Labeled Indications for Human Drugs

Ritu Khare, Jiao Li, & Zhiyong Lu*, Earl Stadtman Investigator

National Center for Biotechnology Information (NCBI)
National Library of Medicine (NLM)
National Institutes of Health (NIH)

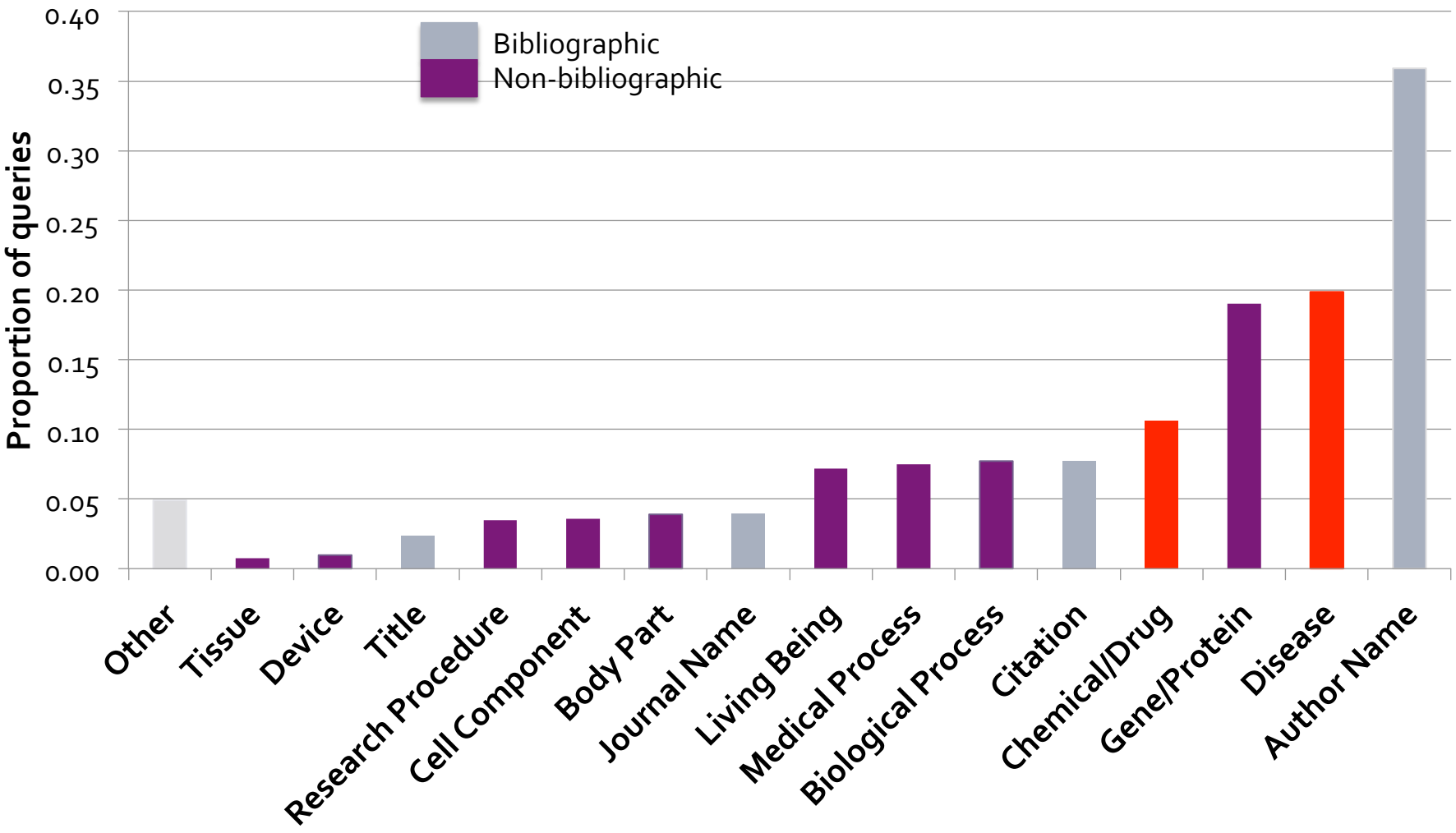
*contact: zhiyong.lu@nih.gov

NCBI

LabeledIn Executive Summary

- Motivation: existing public resources are
  - Not computable
  - In complete
  - Dose-form non-specific
- Method: double-annotation by experts with computer aids for improved efficiency
- Results:
  - 250 popular human drugs in PubMed Health website, e.g.
    - Antidepressant drugs
    - Pain medicine
  - ~8,000 drug and indication pairs are
    - Structured & Normalized
    - Dose-form specific
  - Highly reliable (human inter-annotator agreement is ~95%)
  - Open access
- Future Directions
  - Increase the drug coverage
  - Improve productivity by advanced computer techniques

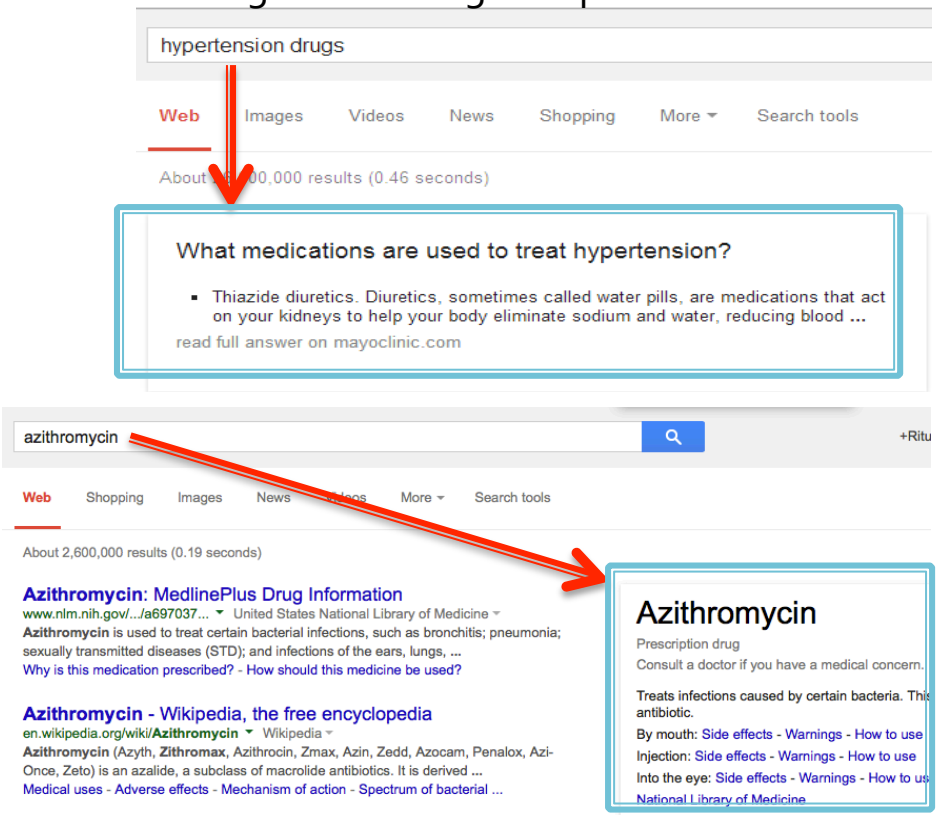
Most searched topics in PubMed: authors; diseases; genes; chemical/Drugs



Dogan et al., Understanding PubMed® user search behavior through log analysis, Database, 2009

Drug-Disease Treatment Relationships

- Question-Answering System
  - Google Knowledge Graph



- Enriching hyperlinks in online health resources
  - Cross-linking Drug and Disease Monographs
- Clinical decision support
  - Medication and Nomenclature Errors (Khare et al. 2013)
  - Recommend Medications
- Computational Methods
  - Drug Repurposing (Li and Lu 2012)
  - Identification of Side Effects

Existing Resources on Drug-Disease Relationships
Not Computable

MedlinePlus®
Trusted Health Information for You

MedicineNet.com
We Bring Doctors' Knowledge to You

Daily Med

PubMed Health

DRUGBANK
Open Data Drug & Drug Target Database

Uses of This Medicine

Diltiazem is used alone or together with other medicines to treat severe chest pain (angina) or high blood pressure (hypertension). High blood pressure adds to the workload of the heart and arteries. If it continues for a long time, the heart and arteries may not function properly. This can damage the blood vessels of the brain, heart, and kidneys, resulting in a stroke, heart failure, or kidney failure. High blood pressure may also increase the risk of heart attacks. These problems may be less likely to occur if blood pressure is controlled .

NOT STRUCTURED

Freebase

Used To Treat /medicine

Used To Treat
Gonorrhea
Chancroid
Chlamydia infection
Staphylococcal infection
Otitis media
Haemophilus Infections

NOT NORMALIZED

National Drug File - Reference Terminology

- Incomplete (Barriere et al. 2011)
- No Differentiation of Dose Forms
  - Ketorolac Drug
    - Injection for pain
    - Ophthalmic Solution for conjunctivitis

NOT PRECISE

Create a computable resource on drug indications based on SPLs

Source: FDA Drug Labels (from DailyMed)

PERPHENAZINE tablet, film coated
[Sandoz Inc]

INDICATIONS AND USAGE

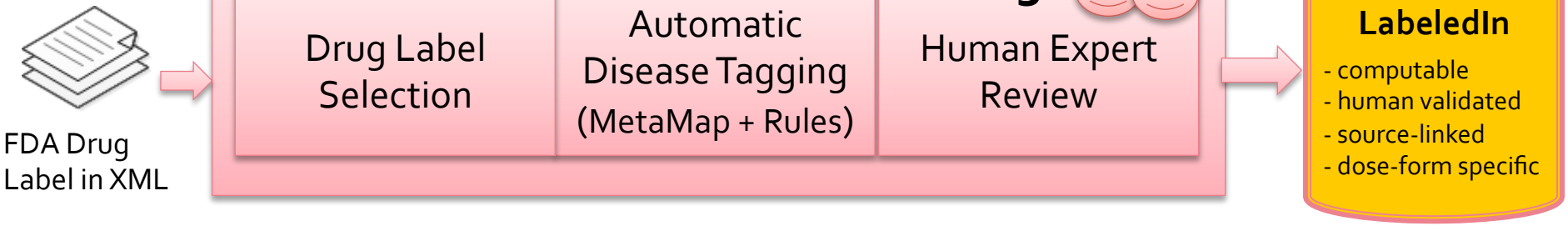
Perphenazine is indicated for use in the treatment of schizophrenia and for the control of severe nausea and vomiting in adults.
Perphenazine has not been shown effective for the management of behavioral complications in patients with mental retardation.

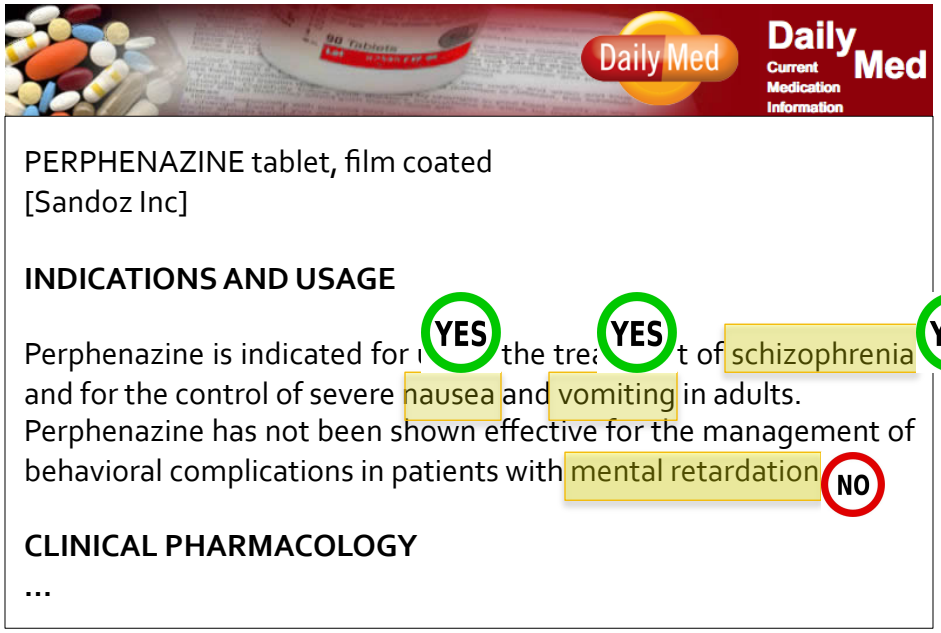
CLINICAL PHARMACOLOGY
…

Related Work

- Automatic approaches:
  - Filter using side effects (SIDER 2, Kuhn et al. 2010)
  - Frequency of occurrence in other drug resources (MEDI, Wei et al. 2013)
  - NLP approach (SPL-X, Kung et al., 2013)
  - Limited in precision and recall
- Manual Annotations
  - *LabeledIn* (Khare, et al., 2014): an expert validated computable resource on drug indications
  - Reliable but expensive to scale

Text Mining + Human Review = LabeledIn





Khare, Li, & Lu: **LabeledIn: Cataloging Labeled Indications for Human Drugs**, JBI, 2014

Human Review Process

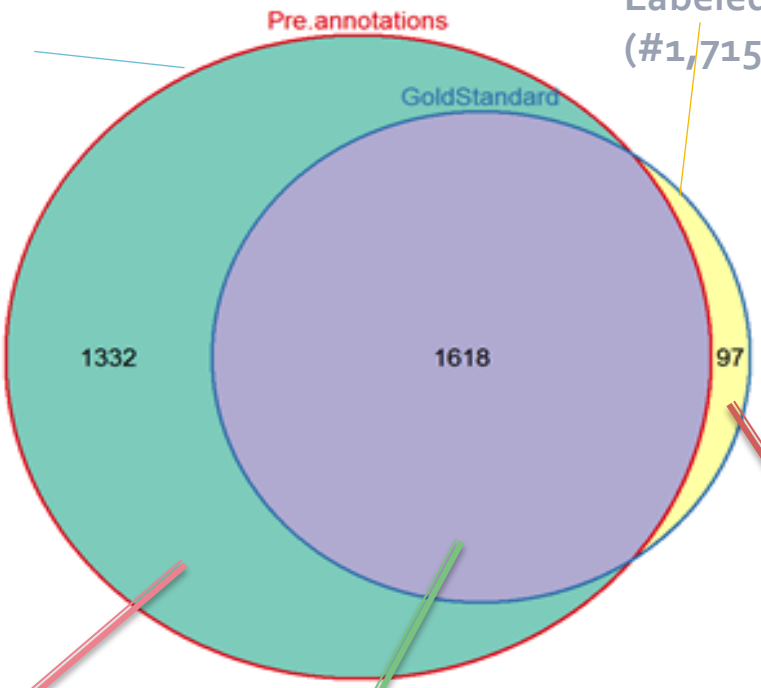
- 3 human reviewers (knowledgeable in biomedicine and indexing)
- Two rounds of review
  - Round #1: Annotating independently
  - Round #2: Resolving disagreements
- Annotation guidelines (e.g. NOT to include contraindications or side effects)
- Pilot study of 100 drug labels for practicing as well as developing guidelines
- Inter-annotator agreements (IAAs)

	Round1	Round1+Round2
Absolute Agreement	89%	94%
Kappa Agreement	77%	88%

Text-mined pre-annotations vs. Human-reviewed results

Pre-annotations (#2950)

LabeledIn (#1,715)



Rejected

Accepted

Human-adjusted/added

Automatic indication detection Accuracies:

Precision	Recall	F_1 Measure
0.55	0.94	0.69

STROMECTOL is indicated for the treatment of Strongyloidiasis of the intestinal tract

STROMECTOL is indicated for the treatment of Onchocerciasis due to nematode parasite. This indication is based on randomized double blind, placebo controlled and comparative studies...

Results: LabeledIn

- Current size
  - 250 ingredients
  - 8,151 drug labels
  - 1,513 drugs <ingredient, dose form, strength> & 7,805 drug-disease pairs
- Features
  - Open access
  - Computable
  - Human-validated
  - Source-linked (original label + human annotations)
- Applications:
  - Computational drug discovery (e.g. re-purposing)
  - Improving existing resources (e.g. SIDER)
  - Decision Support in EMRs (Mayo Clinic)
  - Enabling automatic machine-learning methods
  - Enabling crowdsourcing experiments

Khare, Li, & Lu: LabeledIn: Cataloging Labeled Indications for Human Drugs, JBI 2014

LabeledIn vs SIDER 2 Indications

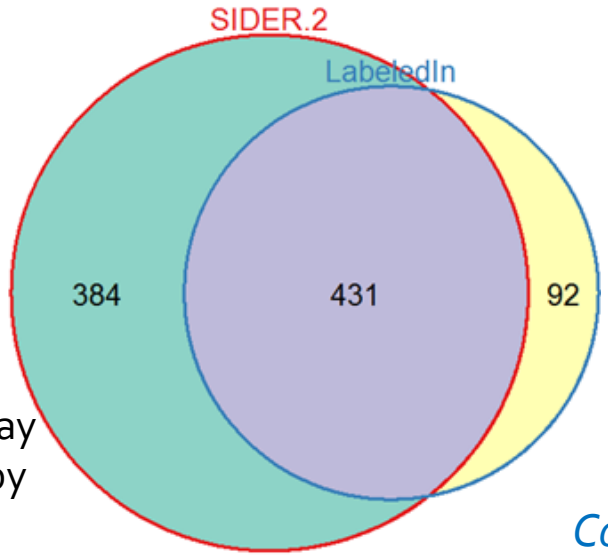
(Systematic Comparison of 50 Drug Labels)

Imprecise Mentions

myocardial infarction

More Specific

partial onset seizures



Unrelated Diseases

...are indicated in the treatment of symptomatic parkinsonism which may follow injury to the nervous system by carbon monoxide intoxication

Disjoint Mentions

biliary and renal colic

Comprehensive Disease Definition

Zollinger-Ellison syndrome

Non-Disease Mentions

Aspirin may be continued, ...use of NSAIDs including salicylates has not been fully explored (see PRECAUTIONS , Drug Interactions)

SIDER 2
Side Effect Resource

Limitations & Future Directions

- To include more drugs into LabeledIn
- To further improve manual curation efficiency by using advanced computer techniques
  - Machine learning
  - Crowdsourcing
  - etc
- To link some general indications to specific procedure/ conditions (e.g. “nausea” and “vomiting” due to “cancer chemotherapy”)
- To improve automatic disease tagging quality in step 2

Acknowledgements

- Three Human indexers
- Pharmacists from Massachusetts college of pharmacy and health sciences
- MetaMap authors - Alan R. Aronson et al.,
- DailyMed developers
- Intramural Research Program of the NIH, National Library of Medicine

Questions?

NCBI

Thank you!

zhiyong.lu@nih.gov

LabeledIn freely available at: ftp.ncbi.nlm.nih.gov/pub/lu/LabeledIn/

NIH Intramural Research Program
Our Research Changes Lives